Optimized Compression for Earth Science Data Using JPEG 2000

Janet C. Rountree[a], Brian N. Webb[a], Thomas J. Flohr[a], Michael W. Marcellin[b]
[a]Science Applications International Corporation, 101 N. Wilmot Suite 400, Tucson, AZ 85711
[b]Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ 85721

***Abstract*-JPEG 2000, a wavelet-based algorithm, is being promulgated by the International Standards Organization (ISO) as the next industrial standard for image data compression. It is both more efficient and more flexible than its predecessor, JPEG. It performs both lossless and lossy compression at a user-selectable compression ratio. Under a previous contract with NASA's Explorer Technology program, SAIC developed a "scan-based" (low-memory) implementation of JPEG 2000 Part 1, the most basic form of the algorithm. JPEG 2000 Part 2, which is currently being finalized by ISO, contains many features of particular interest for Earth Science applications. These include special wavelet filters and decomposition trees for SAR data; single-sample overlap wavelets for artifact reduction; trellis-coded quantization for highest visual quality; and multiple component decorrelation for hyperspectral data. In this project, we will incorporate these Part 2 features into the scan-based implementation of JPEG 2000. The project will continue by testing the optimized software on Earth Science data in a laboratory environment. The software will then be ported to a flight simulation environment and tested there.**

1. INTRODUCTION

JPEG 2000 is the emerging International Standard for digital image compression. It provides superior image quality to the baseline JPEG standard, especially at high compression ratios, and contains many special features that facilitate its adaptation to particular types of imagery.

Fig. 1 is a flow diagram of the JPEG 2000 algorithm. The component transform is used for three-color or for multi-spectral/hyperspectral data, to perform de-correlation in the wavelength dimension. The wavelet transform performs de-correlation in the two spatial dimensions. The quantizer is the principal source of “lossiness” in compression, while the entropy coder is lossless. Finally, the rate controller ensures that the desired compression ratio is reached.

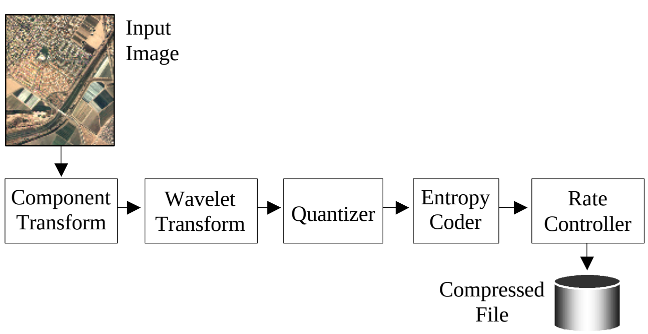

Fig. 1. Flow diagram of the JPEG 2000 encoder.

This work was partially supported by NASA contract NAS5-00186.

Early in the development of the JPEG 2000 standard, a decision was made to divide the technology into Part 1 (Core Coding System) and Part 2 (Extensions). Part 1 contains the features that all decoders must support, in order to be called JPEG 2000 compliant. These include (9x7) and (5x3) wavelet filters with a Mallat decomposition tree, scalar quantization, and three-component color space transforms. There was also a requirement that all technologies accepted for Part 1 would be offered by their originators on a royalty-free, non-discriminatory basis. Technologies that were considered too complex, too limited in their application, or potentially subject to license fees, were placed in Part 2. Unlike Part 1, the Part 2 technologies do not have to be supported as a group by all decoders. One or more Part 2 technologies may be added to a Part 1 decoder to make it Part 2 compliant.

II. THE SCAN-BASED MODE

For technology development purposes, the JPEG 2000 algorithm is embodied in the Verification Model (VM) software, which is maintained by Science Applications International Corporation (SAIC) and the University of Arizona (UA). Early versions of the VM required the entire image to be retained in memory during computation of the compressed file. Later versions required only that the entire image be buffered in the compressed domain, in order to achieve effective rate control. This configuration is sometimes referred to as the “frame-based mode.”

Representatives of the remote sensing community pointed out that airborne and satellite-borne instruments have extremely limited memory, owing to size, weight and power constraints. Moreover, many remote sensing instruments are pushbroom scanners, which naturally build up a large image one line at a time. For these applications, it is desirable to have a JPEG 2000 implementation that buffers up the smallest possible number of image lines. This configuration is called the “scan-based mode.” The scan-based mode is based on the use of "scan elements," which may be either image tiles or "precincts." The precinct, a concept unique to JPEG 2000, is an area in the wavelet domain that corresponds to a location in the image domain. The difference between tiles and precincts is illustrated in Fig. 2.

On the basis of two experiments performed by SAIC/UA and the Centre National d’Etudes Spatiales (CNES), SAIC integrated an implementation of the scan-based mode into the VM [1,2]. This implementation, which incorporated only

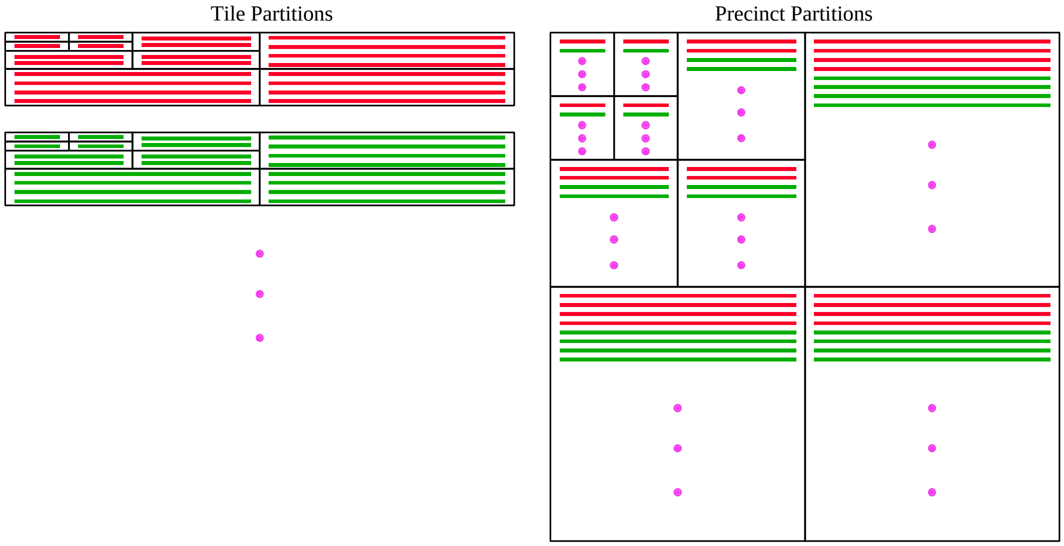

Fig 2. Partitioning an image in tiles and precincts.

Part 1 features, has been described in detail by Flohr et al. [3]. The rate control buffer, which is the largest buffer in the frame-based mode, is set up to contain a user-selectable number of scan elements. A sliding window rate control is then effected by truncating the scan elements in the buffer to achieve the desired bit rate. As a new scan element enters the buffer, bytes are released from the scan element at the head of the window.

III. Part 2 Features In The Scan-based Mode

The technologies included in JPEG 2000 Part 2 are primarily intended for certain niche markets. Many of them offer significant advantages for Earth Science sensors. When combined with the scan-based mode, these Part 2 technologies will constitute an optimized compression algorithm for Earth Science data.

A. The Wavelet Transform

Whereas Part 1 allows only two wavelet filters and one decomposition tree, the user can specify any arbitrary wavelet in Part 2. Experience has shown that for synthetic aperture (SAR) data, improved visual quality can often be obtained by using a longer filter and a more detailed decomposition tree [4,5]. One such decomposition, the packet decomposition, is compared with the standard 5-level Mallat in Fig. 3.

A second wavelet feature of interest in remote sensing is the use of the single sample overlap discrete wavelet transform (SSODWT) [6] or, alternatively, the "odd tile/low pass first" convention (OTLPF) [7] to reduce boundary artifacts at tile edges. Although precincts generally give better image quality than tiles in the scan-based mode [1], they do allow limited error propagation between scan elements. Because of the continuity of the wavelet transform, a bit error in one precinct will cause lower-amplitude errors in neighboring precincts, according to the formula

$$e_l = 2e_{l-1} + k - 2 \qquad (1)$$

where e_l is the extent of errors in level l, $e_{(l-1)}$ is the extent of errors in the previous level, and k is the length of the longest synthesis filter. Fig. 4 gives an example of this error propagation.

Thus if error containment is the primary concern, as it may be in some remote sensing situations where the compressed imagery is to be transmitted over a noisy channel, tiles may be preferred as scan elements despite the possibility of boundary artifacts. Under such circumstances, artifact reduction techniques such as SSODWT and OTLPF may be useful.

The VM implementation of the wavelet transform, including the Part 2 options, is not incompatible with the scan-based mode. However, it currently buffers more than the minimum number of image lines required to complete the sliding window transform. (This minimum is on the order of the maximum vertical filter length.) Some optimization in terms of memory management may be required to obtain the best results for the scan-based mode.

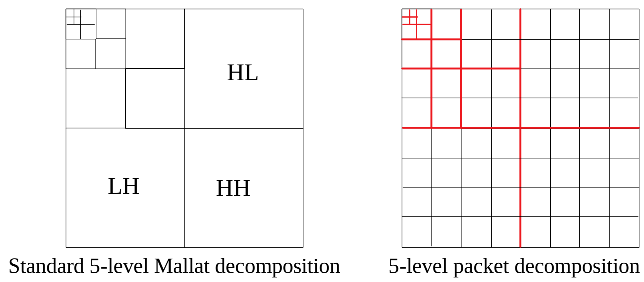

Fig. 3. Comparison of Mallet vs. packet wavelet decomposition structures [5].

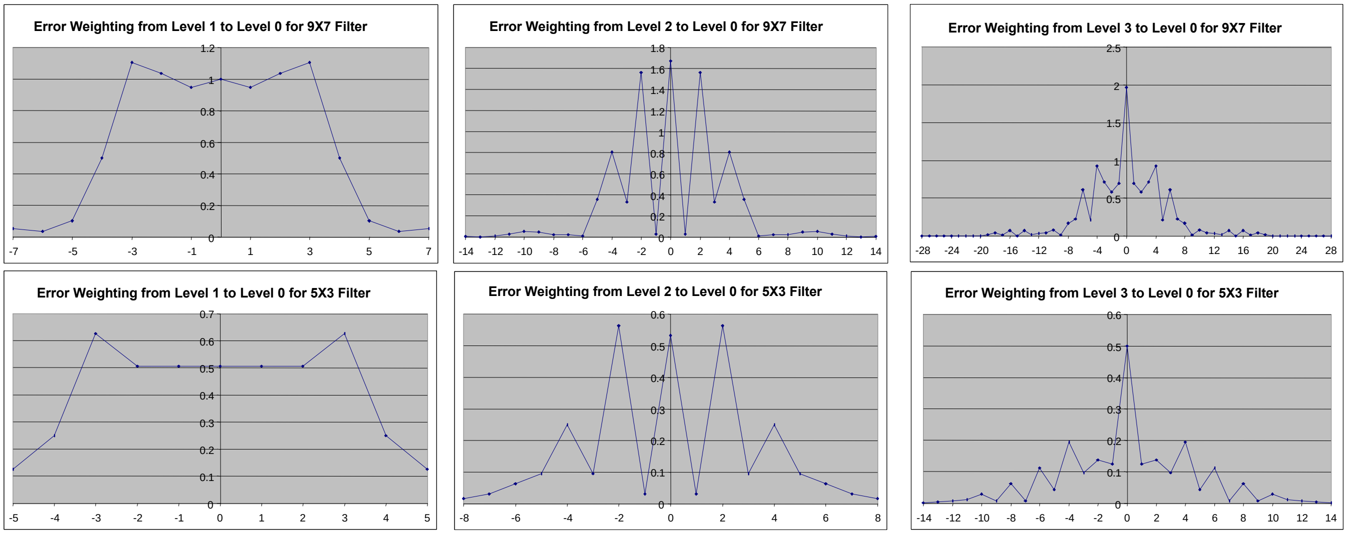

Fig.4. Error propagation as a function of resolution level for two JPEG 2000 filters

B. The multiple component transform

JPEG 2000 Part 1 specifies two transforms, the reversible and irreversible multiple component transforms, that may be applied to the first three components of an image (although as many as 16K components may be present). But multispectral and hyperspectral imagery play a large part in remote sensing science, and these multi-component images are highly correlated in the third (wavelength) dimension. JPEG 2000 Part 2 allows two types of multiple component transform: an arbitrary linear transform (including the Karhuenen-Loeve [KLT] and Differential Pulse Coded Modulation [DPCM] transforms) and a wavelet transform in the third dimension, which is performed independently of the two-dimensional spatial wavelet transform.

The scan-based mode takes advantage of one of the five progression orders allowed in JPEG 2000, namely progression by location, to transmit an image a few lines at a time. The second term in this order is progression by component, so that in fact all the components of a scan element will be output before the encoder begins on the next scan element. Thus it is possible to perform a multiple component transform within a single scan element (even if the scan element is a precinct). However, transforms that require the collection of statistics over the whole image – like the KLT – are clearly ruled out. Other linear transforms, and the third-dimension wavelet transform, are compatible with the scan-based mode. As in the case of the two-dimensional wavelet, some optimization may be required.

C. Trellis-coded quantization (TCQ)

Trellis-coded quantization (TCQ) may be thought of as time-varying scalar quantization, or as an approach to vector quantization [8]. It has been shown to produce better visual quality than scalar quantization [9], especially for detected SAR imagery [10]. So despite its increased complexity, TCQ may be desirable for some remote sensing applications. In the frame-based mode, the step sizes for TCQ are determined by a Lagrangian rate allocator (LRA), which models the statistics for the entire image. The LRA may be used in a single pass, but better results are obtained when the rate allocator is allowed to iterate until it achieves the target bit rate (within a user-selectable tolerance).

In the scan-based mode, iterated rate control is unacceptable because of the need for maximum throughput. In order to achieve effective single-pass rate control, it is necessary to compute the quantization step sizes separately for each scan element. If precincts are used as scan elements, rather than tiles, this procedure is known as “precinct-dependent quantization.” Unlike the scan-based implementation of Part 1 [3], the rate control buffer for Part 2 is part of the quantization object, if explicit quantization is being used. This buffer holds only one scan element at a time (Fig. 5). It is applicable to various forms of scalar quantization, as well as to TCQ.

In our implementation, the LRA collects statistics for the first scan element in the image. For this first scan element, S_1, the "target rate", R_1, of the LRA is set equal to T_1, the desired global rate for the image as a whole. The rate actually achieved after compression of S_1 is A_1. (R_1, T_1, and A_1 are measured in bits per pixel (bpp).) Let D_1 be the size of the initial input file and B_1 be the size of the desired output file after compression. (D_1 is measured in pixels and B_1 is measured in bits.)

For subsequent scan elements, the target rate is modified, based on performance on the preceding scan elements. Thus, for the second scan element, S_2, we set

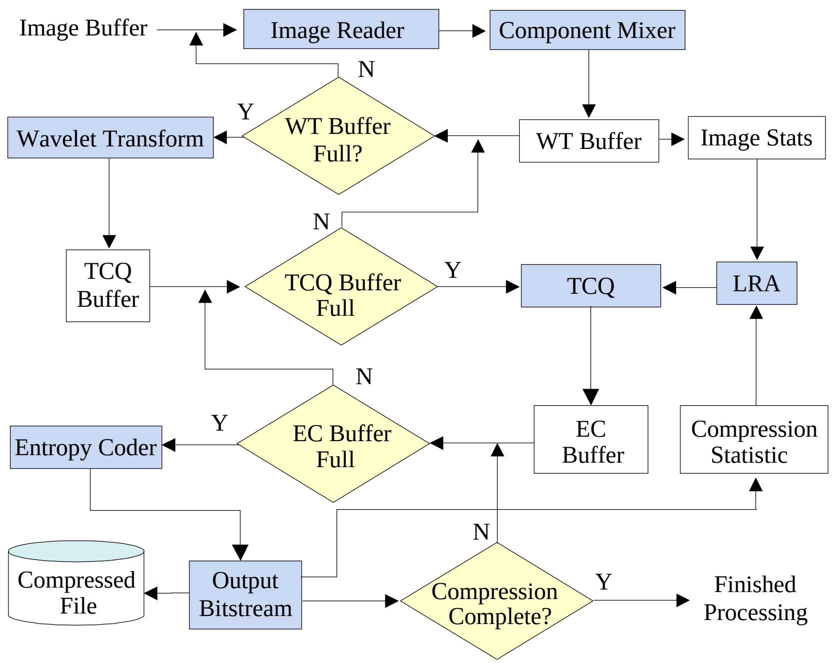

Fig. 5. Scan-based TCQ flow diagram.

$$T_2 = \frac{B_2}{D_2} \tag{2}$$

where D_2 is the input file size after removal of S_1, and B_2 is the remaining space in the output file after compression of S_1. Then the target rate becomes

$$R_2 = T_2 \frac{R_1}{A_1} . \tag{3}$$

Statistics are then gathered for S_2 and LRA is performed using a target rate of R_2. More generally, for the n^{th} scan element,

$$T_n = \frac{B_n}{D_n} \tag{4}$$

$$R_n = T_n \frac{R_{n-1}}{A_{n-1}} \tag{5}$$

where B_n and D_n are the remaining input and output file sizes, respectively, after compression of scan element n-1. LRA is performed using statistics from the n^{th} scan element for a target rate of R_n.

We are also experimenting with the introduction of a damping term to limit the fluctuations of R_n, such that

$$\alpha R_{n-1} \le R_n \le \beta R_{n-1} \tag{6}$$

where $\beta > \alpha$.

This scan-based implementation of TCQ was tested on the four remote sensing images from the JPEG 2000 test set: aerial1 (cropped to 5K x 5K), aerial2, sar1, and sar2. PSNR for the scan-based TCQ implementation was compared with PSNR for frame-based TCQ. As described above, there was no iteration in the rate control for the scan-based mode, while the LRA in the frame-based mode was allowed to iterate. Table 1 shows the difference in PSNR between the single-pass scan-based mode and the iterated frame-based mode, averaged over four images, as a function of bit rate. The performance difference is very small.

TABLE 1

PERFORMANCE DIFFERENCE BETWEEN FULL TCQ AND SCAN-BASED TCQ FOR FOUR REMOTE SENSING IMAGES

Performance difference between Full TCQ and Scan-Based TCQ with 64 High Scan Elements.	
Rate (BPP)	ΔPSNR (dB)
2.0000	– 0.15
1.0000	– 0.12
0.5000	– 0.10
0.2500	– 0.14
0.1250	– 0.19
0.0625	– 0.30

Results for aerial1 are shown graphically in Fig. 6.

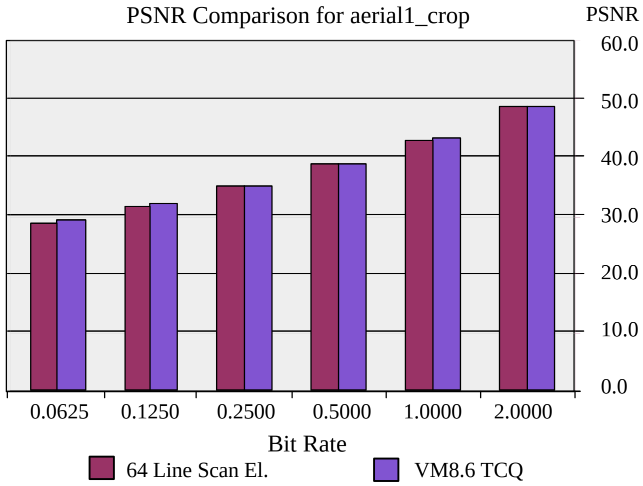

Fig. 6. PSNR Comparison for aerial1

Although our current implementation demonstrates the capability of scan-based TCQ, it is not entirely appropriate for on-board use with a pushbroom scanning sensor. We have made use of the input file size and desired output file size of the image as a whole, which would not be available in a pushbroom scanner, where the number of image lines to be compressed is usually not known at the outset. We plan to modify our approach in the near future to eliminate the use of global input and output file sizes in the rate allocator.

IV. Progression Orders

There are four basic dimensions of progression in the JPEG 2000 bitstream: resolution, quality, spatial location, and component. Different types of progression are achieved by the ordering of packets in the bitstream. Each packet is associated with one component, one (quality) layer, one resolution level, and one location. A bitstream for any desired progression order can be constructed by writing the packets using four nested loops, where any one of the four dimensions can be used as the outermost loop. (The relative order of the remaining loops is generally fixed in the standard.)

In the scan-based mode, the only possible progression order is by spatial location and, as we have seen, the second dimension in this progression is by component. This is the order that will be used in transmitting imagery from a satellite with a pushbroom sensor, after on-board compression. However, once the compressed data are received and archived on the ground, it is possible to re-order the bitstream so as to achieve different compression orders for different clients. This can be done because the coded data within packets are identical regardless of the progression order chosen.

The bitstream contains markers that identify the progression type. Other markers may be written to store the length of every packet in the bitstream. To change the progression order of a bitstream, an application called a parser can read all the markers, change the type of progression in the markers, and write the lengths of the packets out in the new order. Then the packets themselves can be written out in the new order – all without decoding [11].

Thus a bitstream that was received as progression by location can be converted to progression by resolution, and only the lowest resolution sent out to a client who wants only a “thumbnail sketch” of a large number of images. Or the initial bitstream may be converted to progression by quality, after which it can be truncated to produce a lower quality image at a higher compression ratio. Finally, a multi-component image can be reordered so that the most interesting wavelengths are transmitted first.

Our project will include writing a parser for the Earth Science optimized version of JPEG 2000. This will demonstrate the algorithm’s usefulness for archiving and distribution to users, as well as for on-board transmission.

V. Conclusions and Future Plans

It has been shown that several features in JPEG 2000 Part 2 are of potential benefit for Earth Science applications. We are in the process of implementing these features in the scan-based (low-memory) mode, for use on board satellites with pushbroom sensors. Results have been obtained for the scan-based version of the trellis-coded quantizer (TCQ). There is very little performance loss as compared with the fully iterated, frame-based version.

Upon completing the scan-based implementation of the JPEG 2000 Part 2 features described here, we plan to port our software to a flight simulation environment where it can be demonstrated under realistic conditions. It is hoped that this exercise will hasten the day when JPEG 2000 will come into use in satellite-borne applications. We also plan to write a parser that can change the progression order of a JPEG 2000 compressed file, in order to demonstrate the algorithm’s utility for archiving and dissemination.

References

[1] J. C. Rountree, T. J. Flohr, M. W. Marcellin, “Report on core experiment codeff5: scan-based processing mode,” ISO/IEC JTC1/SC29/WG1 N1594, March, 2000.

[2] C. Lambert-Nebout, F. Pellau, “Report on core experiment codeff5: scan-based processing mode,” ISO/IEC JTC1/SC29/WG1 N1581, March, 2000.

[3] T. J. Flohr, M. W. Marcellin, J. C. Rountree, "Scan-based processing with JPEG-2000*," Proc. of SPIE, Appl. of Digital Image Proc.*, pp 347-355, July, 2000.

[4] M. A. Lepley, T. S. Wilkinson, "Wavelet filters for SAR image compression," ISO/IEC JTC1/SC29/WG1 N921, July, 1998.

[5] A. T. Chien, K. S. Miettinen, A. Lan, M. A. Lepley, "Compression of detected SAR imagery*," Proc. of SPIE, Appl. of Digital Image Proc.*, pp 510-520, July, 2000.

[6] B. Berthelot, "Report on core experiment codeff11: reduced overlap in SSWT," ISO/IEC JTC1/SC29/WG1 N1189, March, 1999.

[7] J. Wei, M. R. Pickering, M. R. Frater, J. F. Arnold, J. Boman, W. Zeng, "Boundary artifact reduction using odd tile lengths and the low pass first convention (OTLPF)," *Proc. of SPIE, Appl. of Digital Image Proc.*, July, 2001.

[8] M. W. Marcellin, T. R. Fischer, "Trellis coded quantization of memoryless and Gauss-Markov sources*," IEEE Trans. Commun.,* vol. 38, no. 1, pp 82-93, January, 1990.

[9] T. Chinen, T. J. Flohr, M. W. Marcellin, "Visual assessment of TCQ and Langrangian rate allocation," ISO/IEC JTC1/SC29/WG1 N1307, June, 1999.

[10] M. A. Lepley, A. T. Chien, "Visual testing on SAR," ISO/IEC JTC1/SC29/WG1 N1458, November, 1999.

[11] M. W. Marcellin, M. J. Gormish, A. Bilgin, M. P. Boliek, "An Overview of JPEG-2000," *Proc. of IEEE Data Compression Conference,* Snowbird, Utah, pp 523-541, March, 2000.